

## AXMI-004, A DELTA-ENDOTOXIN GENE AND METHODS FOR ITS USE

### CROSS REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application Serial No. 60/448,810, filed February 20, 2003, the contents of which are herein incorporated by reference in their entirety.

### FIELD OF THE INVENTION

This invention relates to the field of molecular biology. Provided are novel genes that encode pesticidal proteins. These proteins and the nucleic acid sequences that  
5 encode them are useful in preparing pesticidal formulations and in the production of transgenic pest-resistant plants.

### BACKGROUND OF THE INVENTION

*Bacillus thuringiensis* is a Gram-positive spore forming soil bacterium  
10 characterized by its ability to produce crystalline inclusions that are specifically toxic to certain orders and species of insects, but are harmless to plants and other non-targeted organisms. For this reason, compositions including *Bacillus thuringiensis* strains or their insecticidal proteins can be used as environmentally acceptable insecticides to control agricultural insect pests or insect vectors for a variety of human or animal diseases.

15 Crystal (Cry) proteins (delta-endotoxins) from *Bacillus thuringiensis* have potent insecticidal activity against predominantly Lepidopteran, Dipteran, and Coleopteran larvae. These proteins also have shown activity against Hymenoptera, Homoptera, Phthiraptera, Mallophaga, and Acari pest orders, as well as other invertebrate orders such as Nematelminthes, Platyhelminthes, and Sarcomastigophora (Feitelson (1993) The  
20 *Bacillus Thuringiensis* family tree. In *Advanced Engineered Pesticides*. Marcel Dekker, Inc., New York, N.Y.) These proteins were originally classified as CryI to CryV based

primarily on their insecticidal activity. The major classes were Lepidoptera-specific (I), Lepidoptera- and Diptera-specific (II), Coleoptera-specific (III), Diptera-specific (IV), and nematode-specific (V) and (VI). The proteins were further classified into subfamilies; more highly related proteins within each family were assigned divisional letters such as Cry1A, Cry1B, Cry1C, etc. Even more closely related proteins within each division were given names such as Cry1C1, Cry1C2, etc.

A new nomenclature was recently described for the Cry genes based upon amino acid sequence homology rather than insect target specificity (Crickmore *et al.* (1998) *Microbiol. Mol. Biol. Rev.* 62:807-813). In the new classification, each toxin is assigned a unique name incorporating a primary rank (an Arabic number), a secondary rank (an uppercase letter), a tertiary rank (a lowercase letter), and a quaternary rank (another Arabic number). In the new classification, Roman numerals have been exchanged for Arabic numerals in the primary rank. Proteins with less than 45% sequence identity have different primary ranks, and the criteria for secondary and tertiary ranks are 78% and 95%, respectively.

The crystal protein does not exhibit insecticidal activity until it has been ingested and solubilized in the insect midgut. The ingested protoxin is hydrolyzed by proteases in the insect digestive tract to an active toxic molecule. (Höfte and Whiteley (1989) *Microbiol. Rev.* 53:242-255). This toxin binds to apical brush border receptors in the midgut of the target larvae and inserts into the apical membrane creating ion channels or pores, resulting in larval death.

Delta-endotoxins generally have five conserved sequence domains, and three conserved structural domains (see, for example, de Maagd *et al.* (2001) *Trends Genetics* 17:193-199). The first conserved structural domain consists of seven alpha helices and is involved in membrane insertion and pore formation. Domain II consists of three beta-sheets arranged in a Greek key configuration, and domain III consists of two antiparallel beta-sheets in 'jelly-roll' formation (de Maagd *et al.* (2001) *supra*). Domains II and III are involved in receptor recognition and binding, and are therefore considered determinants of toxin specificity.

Because of the devastation that insects can confer, there is a continual need to discover new forms of *Bacillus thuringiensis* delta-endotoxins.

## SUMMARY OF INVENTION

Compositions and methods for conferring pesticide resistance to bacteria, plants, plant cells, tissues, and seeds are provided. Compositions include isolated nucleic acid molecules encoding sequences for delta-endotoxin polypeptides, vectors comprising  
5 those nucleic acid molecules, and host cells comprising the vectors. Compositions also include isolated or recombinant polypeptide sequences of the endotoxin, compositions comprising these polypeptides, and antibodies to those polypeptides. The nucleotide sequences can be used in DNA constructs or expression cassettes for transformation and expression in organisms, including microorganisms and plants. The nucleotide or amino  
10 acid sequences may be synthetic sequences that have been designed for optimum expression. Compositions also comprise transformed bacteria, plants, plant cells, tissues, and seeds.

In particular, the present invention provides for isolated nucleic acid molecules comprising the nucleotide sequences encoding the amino acid sequences shown in SEQ  
15 ID NOS:3 and 5 and the nucleotide sequences set forth in SEQ ID NOS:1, 2, and 4, as well as variants and fragments thereof. Nucleotide sequences that are complementary to a nucleotide sequence of the invention, or that hybridize to a sequence of the invention are also encompassed.

Methods are provided for producing the polypeptides of the invention, and for  
20 using those polypeptides for controlling or killing a lepidopteran or coleopteran pest.

The compositions and methods of the invention are useful for the production of organisms with pesticide resistance, specifically bacteria and plants. These organisms and compositions derived from them are desirable for agricultural purposes. The compositions of the invention are also useful for generating altered or improved delta-  
25 endotoxin proteins that have pesticidal activity, or for detecting the presence of delta-endotoxin proteins or nucleic acids in products or organisms.

## DESCRIPTION OF FIGURES

Figure 1 shows an alignment of AXMI-004 (SEQ ID NO:3) with cry1Ac (SEQ ID  
30 NO:6), cry1Ca (SEQ ID NO:7), cry2Aa (SEQ ID NO:8), cry3Aa1 (SEQ ID NO:9), cry1Ia (SEQ ID NO:10), and cry7Aa (SEQ ID NO:11). Toxins having C-terminal non-

toxic domains were artificially truncated as shown. The alignment shows the most highly conserved amino acid residues highlighted in black, and highly conserved amino acid residues highlighted in gray. Conserved group 1 is found from about amino acid residue 174 to about 196 of SEQ ID NO:3. Conserved group 2 is found from about amino acid residue 250 to about 292 of SEQ ID NO:3. Conserved group 3 is found from about amino acid residue 476 to about 521 of SEQ ID NO:3. Conserved group 4 is found from about amino acid residue 542 to about 552 of SEQ ID NO:3. Conserved group 5 is found from about amino acid residue 618 to about 628 of SEQ ID NO:3

Figure 2 shows a photograph of a 4-20% gradient SDS acrylamide gel. Lanes 1-4 contain various concentrations of sporulated *Bacillus* cell culture expressing 69 kD AXMI-004 protein. Lanes 5-8 contain various concentrations of BSA. Lane 9 contains a size marker. An arrow indicates the 69 kD band.

#### DETAILED DESCRIPTION

The present invention is drawn to compositions and methods for regulating pest resistance in organisms, particularly plants or plant cells. The methods involve transforming organisms with a nucleotide sequence encoding a delta-endotoxin protein of the invention. In particular, the nucleotide sequences of the invention are useful for preparing plants and microorganisms that possess pesticidal activity. Thus, transformed bacteria, plants, plant cells, plant tissues and seeds are provided. Compositions are delta-endotoxin nucleic acids and proteins of *Bacillus thuringiensis*. The sequences find use in the construction of expression vectors for subsequent transformation into organisms of interest, as probes for the isolation of other delta-endotoxin genes, and for the generation of altered pesticidal proteins by methods known in the art, such as domain swapping or DNA shuffling. The proteins find use in controlling or killing lepidopteran or coleopteran pest populations and for producing compositions with pesticidal activity.

#### Definitions

By “delta-endotoxin” is intended a toxin from *Bacillus thuringiensis* that has toxic activity against one or more pests, including, but not limited to, members of the

Lepidoptera, Diptera, and Coleoptera orders. In some cases, delta-endotoxin proteins have been isolated from other organisms, including *Clostridium bifermentans* and *Paenibacillus popilliae*. Delta-endotoxin proteins include amino acid sequences deduced from the full-length nucleotide sequences disclosed herein, and amino acid sequences that are shorter than the full-length sequences, either due to the use of an alternate downstream start site, or due to processing that produces a shorter protein having pesticidal activity. Processing may occur in the organism the protein is expressed in, or in the pest after ingestion of the protein. Delta-endotoxins include proteins identified as cry1 through cry43, cyt1 and cyt2, and Cyt-like toxin. There are currently over 250 known species of delta-endotoxins with a wide range of specificities and toxicities. For an expansive list see Crickmore *et al.* (1998), *Microbiol. Mol. Biol. Rev.* 62:807-813, and for regular updates see Crickmore *et al.* (2003) "Bacillus thuringiensis toxin nomenclature," at [www.biols.susx.ac.uk/Home/Neil\\_Crickmore/Bt/index](http://www.biols.susx.ac.uk/Home/Neil_Crickmore/Bt/index).

Bacterial genes, such as the AXMI-004 gene of this invention, quite often possess multiple methionine initiation codons in proximity to the start of the open reading frame. Often, translation initiation at one or more of these start codons will lead to generation of a functional protein. These start codons can include ATG codons. However, bacteria such as *Bacillus sp.* also recognize the codon GTG as a start codon, and proteins that initiate translation at GTG codons contain a methionine at the first amino acid. Furthermore, it is not often determined *a priori* which of these codons are used naturally in the bacterium. Thus, it is understood that use of one of the alternate methionine codons may also lead to generation of delta-endotoxin proteins that encode pesticidal activity. For example, an alternate start site for a delta-endotoxin protein of the invention is at base pair 385 of SEQ ID NO:1. Translation from this alternate start site results in the amino acid sequence found in SEQ ID NO:5. These delta-endotoxin proteins are encompassed in the present invention and may be used in the methods of the present invention.

By "plant cell" is intended all known forms of plant, including undifferentiated tissue (e.g. callus), suspension culture cells, protoplasts, leaf cells, root cells, phloem cells, plant seeds, pollen, propagules, embryos and the like. By "plant expression cassette" is intended a DNA construct that is capable of resulting in the expression of a protein from an open reading frame in a plant cell. Typically these contain a promoter

and a coding sequence. Often, such constructs will also contain a 3' untranslated region. Such constructs may contain a 'signal sequence' or 'leader sequence' to facilitate co-translational or post-translational transport of the peptide to certain intracellular structures such as the chloroplast (or other plastid), endoplasmic reticulum, or Golgi apparatus.

5 By "signal sequence" is intended a sequence that is known or suspected to result in cotranslational or post-translational peptide transport across the cell membrane. In eukaryotes, this typically involves secretion into the Golgi apparatus, with some resulting glycosylation. By "leader sequence" is intended any sequence that when translated, results in an amino acid sequence sufficient to trigger co-translational transport of the  
10 peptide chain to a sub-cellular organelle. Thus, this includes leader sequences targeting transport and/or glycosylation by passage into the endoplasmic reticulum, passage to vacuoles, plastids including chloroplasts, mitochondria, and the like.

By "plant transformation vector" is intended a DNA molecule that is necessary for efficient transformation of a plant cell. Such a molecule may consist of one or more  
15 plant expression cassettes, and may be organized into more than one 'vector' DNA molecule. For example, binary vectors are plant transformation vectors that utilize two non-contiguous DNA vectors to encode all requisite cis- and trans-acting functions for transformation of plant cells (Hellens and Mullineaux (2000) *Trends in Plant Science* 5:446-451). "Vector" refers to a nucleic acid construct designed for transfer between  
20 different host cells. "Expression vector" refers to a vector that has ability to incorporate, integrate and express heterologous DNA sequences or fragments in a foreign cell.

"Transgenic plants" or "transformed plants" or "stably transformed plants or cells or tissues" refers to plants that have incorporated or integrated exogenous nucleic acid sequences or DNA fragments into the plant cell. These nucleic acid sequences include  
25 those that are exogenous, or not present in the untransformed plant cell, as well as those that may be endogenous, or present in the untransformed plant cell. "Heterologous" generally refers to the nucleic acid sequences that are not endogenous to the cell or part of the native genome in which they are present, and have been added to the cell by infection, transfection, microinjection, electroporation, microprojection, or the like.

30 "Promoter" refers to a nucleic acid sequence that functions to direct transcription of a downstream coding sequence. The promoter together with other transcriptional and

translational regulatory nucleic acid sequences (also termed "control sequences") are necessary for the expression of a DNA sequence of interest.

Provided herein are novel isolated nucleotide sequences that confer pesticidal activity. Also provided are the amino acid sequences for the delta-endotoxin proteins.

5 The protein resulting from translation of this gene allows cells to control or kill pests that ingest it.

An "isolated" or "purified" nucleic acid molecule or protein, or biologically active portion thereof, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. Preferably, an "isolated" nucleic acid is free of sequences (preferably protein encoding sequences) that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For purposes of the invention, "isolated" when used to refer to nucleic acid molecules excludes isolated chromosomes.

10 For example, in various embodiments, the isolated delta-endotoxin-encoding nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequence that naturally flanks the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. A delta-endotoxin protein that is substantially free of cellular material includes preparations of protein having less than

15 about 30%, 20%, 10%, or 5% (by dry weight) of non-delta-endotoxin protein (also referred to herein as a "contaminating protein"). Various aspects of the invention are described in further detail in the following subsections.

20

#### Isolated Nucleic Acid Molecules, and Variants and Fragments Thereof

25 One aspect of the invention pertains to isolated nucleic acid molecules comprising nucleotide sequences encoding delta-endotoxin proteins and polypeptides or biologically active portions thereof, as well as nucleic acid molecules sufficient for use as hybridization probes to identify delta-endotoxin encoding nucleic acids. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or

30 genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA

generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

Nucleotide sequences encoding the proteins of the present invention include the sequences set forth in SEQ ID NOS:1, 2, and 4, and complements thereof. By

5 "complement" is intended a nucleotide sequence that is sufficiently complementary to a given nucleotide sequence such that it can hybridize to the given nucleotide sequence to thereby form a stable duplex. The corresponding amino acid sequences for the delta-endotoxin proteins encoded by these nucleotide sequences are set forth in SEQ ID NOS:3 and 5.

10 Nucleic acid molecules that are fragments of these delta-endotoxin-encoding nucleotide sequences are also encompassed by the present invention. By "fragment" is intended a portion of the nucleotide sequence encoding a delta-endotoxin protein. A fragment of a nucleotide sequence may encode a biologically active portion of a delta-endotoxin protein, or it may be a fragment that can be used as a hybridization probe or  
15 PCR primer using methods disclosed below. Nucleic acid molecules that are fragments of a delta-endotoxin nucleotide sequence comprise at least about 15, 20, 50, 75, 100, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850 nucleotides, or up to the number of nucleotides present in a full-length delta-  
20 endotoxin-encoding nucleotide sequence disclosed herein (for example, 2190 nucleotides for SEQ ID NO:1, 1890 nucleotides for SEQ ID NO:2, and 1806 nucleotides for SEQ ID NO:4) depending upon the intended use. Fragments of the nucleotide sequences of the present invention will encode protein fragments that retain the biological activity of the delta-endotoxin protein and, hence, retain pesticidal activity. By "retains activity" is  
25 intended that the fragment will have at least about 30%, preferably at least about 50%, more preferably at least about 70%, even more preferably at least about 80% of the pesticidal activity of the delta-endotoxin protein. Methods for measuring pesticidal activity are well known in the art. See, for example, Czapla and Lang (1990) *J. Econ. Entomol.* 83(6): 2480-2485; Andrews *et al.* (1988) *Biochem. J.* 252:199-206; Marrone *et al.* (1985) *J. of Economic Entomology* 78:290-293; and U.S. Patent No. 5,743,477, all of  
30 which are herein incorporated by reference in their entirety.



A fragment of a delta-endotoxin encoding nucleotide sequence that encodes a biologically active portion of a protein of the invention will encode at least about 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, or 600 contiguous amino acids, or up to the total number of amino acids present in a full-length delta-endotoxin protein of the invention (for example, 629 amino acids for SEQ ID NO:3 and 601 for SEQ ID NO:5).

Preferred delta-endotoxin proteins of the present invention are encoded by a nucleotide sequence sufficiently identical to a nucleotide sequence of SEQ ID NO:1, 2, or 4. By "sufficiently identical" is intended an amino acid or nucleotide sequence that has at least about 60% or 65% sequence identity, preferably about 70% or 75% sequence identity, more preferably about 80% or 85% sequence identity, most preferably about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity compared to a reference sequence using one of the alignment programs described herein using standard parameters. One of skill in the art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning, and the like.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., percent identity = number of identical positions/total number of positions (e.g., overlapping positions) x 100). In one embodiment, the two sequences are the same length. The percent identity between two sequences can be determined using techniques similar to those described below, with or without allowing gaps. In calculating percent identity, typically exact matches are counted.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A nonlimiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul *et al.* (1990) *J. Mol. Biol.* 215:403.

BLAST nucleotide searches can be performed with the BLASTN program, score = 100, wordlength = 12, to obtain nucleotide sequences homologous to delta-endotoxin nucleic acid molecules of the invention. BLAST protein searches can be performed with the BLASTX program, score = 50, wordlength = 3, to obtain amino acid sequences

5 homologous to delta-endotoxin protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389. Alternatively, PSI-Blast can be used to perform an iterated search that detects distant relationships between molecules. See Altschul *et al.* (1997) *supra*. When utilizing BLAST, Gapped BLAST, and PSI-Blast  
10 programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) can be used. See, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). Another non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the ClustalW algorithm (Higgins *et al.* (1994) *Nucleic Acids Res.* 22:4673-4680). ClustalW compares sequences and aligns the entirety of the amino acid or DNA sequence, and thus can  
15 provide data about the sequence conservation of the entire amino acid sequence. The ClustalW algorithm is used in several commercially available DNA/amino acid analysis software packages, such as the ALIGNX module of the vector NTi Program Suite (Informax, Inc). After alignment of amino acid sequences with ClustalW, the percent amino acid identity can be assessed. A non-limiting example of a software program  
20 useful for analysis of ClustalW alignments is GeneDoc™. Genedoc™ (Karl Nicholas) allows assessment of amino acid (or DNA) similarity and identity between multiple proteins. Another non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0), which is part of  
25 the GCG sequence alignment software package (available from Accelrys, Inc., 9865 Scranton Rd., San Diego, California, USA). When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

A preferred program is GAP version 10, which used the algorithm of Needleman  
30 and Wunsch (1970) *supra*. GAP Version 10 may be used with the following parameters: % identity and % similarity for a nucleotide sequence using GAP Weight of 50 and

Length Weight of 3, and the nwsgapdna.cmp scoring matrix; % identity and % similarity for an amino acid sequence using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 Scoring Matrix. Equivalent programs may also be used. By “equivalent program” is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by GAP Version 10.

The invention also encompasses variant nucleic acid molecules. "Variants" of the delta-endotoxin-encoding nucleotide sequences include those sequences that encode the delta-endotoxin proteins disclosed herein but that differ conservatively because of the degeneracy of the genetic code as well as those that are sufficiently identical as discussed above. Naturally occurring allelic variants can be identified with the use of well-known molecular biology techniques, such as polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant nucleotide sequences also include synthetically derived nucleotide sequences that have been generated, for example, by using site-directed mutagenesis but which still encode the delta-endotoxin proteins disclosed in the present invention as discussed below. Variant proteins encompassed by the present invention are biologically active, that is they continue to possess the desired biological activity of the native protein, that is, retaining pesticidal activity. By “retains activity” is intended that the variant will have at least about 30%, preferably at least about 50%, more preferably at least about 70%, even more preferably at least about 80% of the pesticidal activity of the native protein. Methods for measuring pesticidal activity are well known in the art. See, for example, Czapla and Lang (1990) *J. Econ. Entomol.* 83(6): 2480-2485; Andrews *et al.* (1988) *Biochem. J.* 252:199-206; Marrone *et al.* (1985) *J. of Economic Entomology* 78:290-293; and U.S. Patent No. 5,743,477, all of which are herein incorporated by reference in their entirety.

The skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of the invention thereby leading to changes in the amino acid sequence of the encoded delta-endotoxin proteins, without altering the biological activity of the proteins. Thus, variant isolated nucleic acid molecules can be created by introducing one or more nucleotide substitutions, additions, or deletions into

the corresponding nucleotide sequence disclosed herein, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Such variant nucleotide sequences are also encompassed  
5 by the present invention.

For example, preferably conservative amino acid substitutions may be made at one or more predicted, preferably nonessential amino acid residues. A "nonessential" amino acid residue is a residue that can be altered from the wild-type sequence of a delta-endotoxin protein without altering the biological activity, whereas an "essential" amino  
10 acid residue is required for biological activity. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar  
15 side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine).

There are generally five highly conserved regions among the delta-endotoxin  
20 proteins, concentrated largely in the center of the domain or at the junction between domains (Rajamohan *et al.* (1998) *Prog. Nucleic Acid Res. Mol. Biol.* 60:1 –23). The blocks of conserved amino acids for various delta-endotoxins as well as consensus sequences may be found in Schnepf *et al.* (1998) *Microbio. Mol. Biol. Rev.* 62:775-806 and Lereclus *et al.* (1989) *Role, Structure, and Molecular Organization of the Genes*  
25 *Coding for the Parasporal d-endotoxins of Bacillus thuringiensis. In Regulation of Prokaryotic Development.* Issar Smit, Slepecky, R.A., Setlow, P. American Society for Microbiology, Washington, D.C. 20006. It has been proposed that delta-endotoxins having these conserved regions may share a similar structure, consisting of three domains (Li *et al.* (1991) *Nature* 353: 815-821). Domain I has the highest similarity between  
30 delta-endotoxins (Bravo (1997) *J. Bacteriol.* 179:2793-2801).

Amino acid substitutions may be made in nonconserved regions that retain function. In general, such substitutions would not be made for conserved amino acid residues, or for amino acid residues residing within a conserved motif, where such residues are essential for protein activity. Examples of residues that are conserved and that may be essential for protein activity include, for example, residues 103, 122, 124, 168, 170, 180, 185-188, 190, 191, 196, 204, 205, 225, 257, 262, 287, 321, 496, 508, 509, and 575 of SEQ ID NO:3. Examples of residues that are conserved but that may allow conservative amino acid substitutions and still retain activity include, for example, residues 106, 109, 110, 125, 132, 139, 152, 167, 172, 183, 194, 195, 199, 211, 233, 240, 260, 271, 329, 332, 336, 480, 511, 587, 620, and 626 of SEQ ID NO:3. However, one of skill in the art would understand that functional variants may have minor conserved or nonconserved alterations in the conserved residues.

Alternatively, variant nucleotide sequences can be made by introducing mutations randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for ability to confer pesticidal activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly, and the activity of the protein can be determined using standard assay techniques.

Using methods such as PCR, hybridization, and the like corresponding delta-endotoxin sequences can be identified, such sequences having substantial identity to the sequences of the invention. See, for example, Sambrook J., and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) and Innis, et al. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, NY).

In a hybridization method, all or part of a delta-endotoxin nucleotide sequence can be used to screen cDNA or genomic libraries. Methods for construction of such cDNA and genomic libraries are generally known in the art and are disclosed in Sambrook and Russell, 2001. The so-called hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as  $^{32}\text{P}$ , or any other detectable marker, such as other radioisotopes, a fluorescent compound, an enzyme, or an enzyme co-factor. Probes for

hybridization can be made by labeling synthetic oligonucleotides based on the known delta-endotoxin encoding nucleotide sequence disclosed herein. Degenerate primers designed on the basis of conserved nucleotides or amino acid residues in the nucleotide sequence or encoded amino acid sequence can additionally be used. The probe typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably at least about 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 consecutive nucleotides of a delta-endotoxin-encoding nucleotide sequence of the invention or a fragment or variant thereof. Preparation of probes for hybridization is generally known in the art and is disclosed in Sambrook and Russell, 2001, herein incorporated by reference.

In hybridization techniques, all or part of a known nucleotide sequence is used as a probe that selectively hybridizes to other corresponding nucleotide sequences present in a population of cloned genomic DNA fragments or cDNA fragments (i.e., genomic or cDNA libraries) from a chosen organism. The hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as  $^{32}\text{P}$ , or any other detectable marker. Thus, for example, probes for hybridization can be made by labeling synthetic oligonucleotides based on the delta-endotoxin sequence of the invention. Methods for preparation of probes for hybridization and for construction of cDNA and genomic libraries are generally known in the art and are disclosed in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

For example, an entire delta-endotoxin sequence disclosed herein, or one or more portions thereof, may be used as a probe capable of specifically hybridizing to corresponding delta-endotoxin-like sequences and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique and are preferably at least about 10 nucleotides in length, and most preferably at least about 20 nucleotides in length. Such probes may be used to amplify corresponding delta-endotoxin sequences from a chosen organism by PCR. This technique may be used to isolate additional coding sequences from a desired organism or as a diagnostic assay to determine the presence of coding sequences in an organism. Hybridization techniques

include hybridization screening of plated DNA libraries (either plaques or colonies; see, for example, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Hybridization of such sequences may be carried out under stringent conditions.

5 By "stringent conditions" or "stringent hybridization conditions" is intended conditions under which a probe will hybridize to its target sequence to a detectably greater degree than to other sequences (e.g., at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences that are 100%  
10 complementary to the probe can be identified (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing). Generally, a probe is less than about 1000 nucleotides in length, preferably less than 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less  
15 than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include  
20 hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include  
25 hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C. Optionally, wash buffers may comprise about 0.1% to about 1% SDS. Duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours.

Specificity is typically the function of post-hybridization washes, the critical  
30 factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the  $T_m$  can be approximated from the equation of Meinkoth and Wahl

(1984) *Anal. Biochem.* 138:267-284:  $T_m = 81.5^{\circ}\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61$   
(% form) - 500/L; where M is the molarity of monovalent cations, %GC is the percentage  
of guanosine and cytosine nucleotides in the DNA, % form is the percentage of  
formamide in the hybridization solution, and L is the length of the hybrid in base pairs.

5 The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of a  
complementary target sequence hybridizes to a perfectly matched probe.  $T_m$  is reduced  
by about  $1^{\circ}\text{C}$  for each 1% of mismatching; thus,  $T_m$ , hybridization, and/or wash  
conditions can be adjusted to hybridize to sequences of the desired identity. For example,  
if sequences with  $\geq 90\%$  identity are sought, the  $T_m$  can be decreased  $10^{\circ}\text{C}$ . Generally,  
10 stringent conditions are selected to be about  $5^{\circ}\text{C}$  lower than the thermal melting point  
( $T_m$ ) for the specific sequence and its complement at a defined ionic strength and pH.  
However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3,  
or  $4^{\circ}\text{C}$  lower than the thermal melting point ( $T_m$ ); moderately stringent conditions can  
utilize a hybridization and/or wash at 6, 7, 8, 9, or  $10^{\circ}\text{C}$  lower than the thermal melting  
15 point ( $T_m$ ); low stringency conditions can utilize a hybridization and/or wash at 11, 12,  
13, 14, 15, or  $20^{\circ}\text{C}$  lower than the thermal melting point ( $T_m$ ). Using the equation,  
hybridization and wash compositions, and desired  $T_m$ , those of ordinary skill will  
understand that variations in the stringency of hybridization and/or wash solutions are  
inherently described. If the desired degree of mismatching results in a  $T_m$  of less than  
20  $45^{\circ}\text{C}$  (aqueous solution) or  $32^{\circ}\text{C}$  (formamide solution), it is preferred to increase the SSC  
concentration so that a higher temperature can be used. An extensive guide to the  
hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in*  
*Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I,  
Chapter 2 (Elsevier, New York); and Ausubel *et al.*, eds. (1995) *Current Protocols in*  
25 *Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York).  
See Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold  
Spring Harbor Laboratory Press, Plainview, New York).

#### Isolated Proteins and Variants and Fragments Thereof

30 Delta-endotoxin proteins are also encompassed within the present invention. By  
"delta-endotoxin protein" is intended a protein having the amino acid sequence set forth



in SEQ ID NO:3 or 5. Fragments, biologically active portions, and variants thereof are also provided, and may be used to practice the methods of the present invention.

"Fragments" or "biologically active portions" include polypeptide fragments comprising a portion of an amino acid sequence encoding a delta-endotoxin protein as set forth in SEQ ID NO:3 or 5 and that retains pesticidal activity. A biologically active portion of a delta-endotoxin protein can be a polypeptide that is, for example, 10, 25, 50, 100 or more amino acids in length. Such biologically active portions can be prepared by recombinant techniques and evaluated for pesticidal activity. Methods for measuring pesticidal activity are well known in the art. See, for example, Czapla and Lang (1990) *J. Econ. Entomol.* 83(6): 2480-2485; Andrews *et al.* (1988) *Biochem. J.* 252:199-206; Marrone *et al.* (1985) *J. of Economic Entomology* 78:290-293; and U.S. Patent No. 5,743,477, all of which are herein incorporated by reference in their entirety. As used here, a fragment comprises at least 8 contiguous amino acids of SEQ ID NO:3 or 5. The invention encompasses other fragments, however, such as any fragment in the protein greater than about 10, 20, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, and 600 amino acids.

By "variants" is intended proteins or polypeptides having an amino acid sequence that is at least about 60%, 65%, preferably about 70%, 75%, more preferably about 80%, 85%, most preferably about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to the amino acid sequence of SEQ ID NO:3 or 5. Variants also include polypeptides encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule of SEQ ID NO:1, 2, or 4, or a complement thereof, under stringent conditions. Such variants generally retain pesticidal activity. Variants include polypeptides that differ in amino acid sequence due to mutagenesis. Variant proteins encompassed by the present invention are biologically active, that is they continue to possess the desired biological activity of the native protein, that is, retaining pesticidal activity. Methods for measuring pesticidal activity are well known in the art. See, for example, Czapla and Lang (1990) *J. Econ. Entomol.* 83(6): 2480-2485; Andrews *et al.* (1988) *Biochem. J.* 252:199-206; Marrone *et al.* (1985) *J. of Economic Entomology* 78:290-293; and U.S. Patent No. 5,743,477, all of which are herein incorporated by reference in their entirety.

### Altered or Improved Variants

It is recognized that DNA sequences of a delta-endotoxin may be altered by various methods, and that these alterations may result in DNA sequences encoding proteins with amino acid sequences different than that encoded by the delta-endotoxin of the present invention. This protein may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of the delta-endotoxin protein can be prepared by mutations in the DNA. This may also be accomplished by one of several forms of mutagenesis and/or in directed evolution. In some aspects, the changes encoded in the amino acid sequence will not substantially affect the function of the protein. Such variants will possess the desired pesticidal activity. However, it is understood that the ability of delta-endotoxin to confer pesticidal activity may be improved by the use of such techniques upon the compositions of this invention. For example, one may express delta-endotoxin in host cells that exhibit high rates of base misincorporation during DNA replication, such as XL-1 Red (Stratagene). After propagation in such strains, one can isolate the delta-endotoxin DNA (for example by preparing plasmid DNA, or by amplifying by PCR and cloning the resulting PCR fragment into a vector), culture the delta-endotoxin mutations in a non-mutagenic strain, and identify mutated delta-endotoxin genes with pesticidal activity, for example by performing an assay to test for pesticidal activity. Generally, the protein is mixed and used in feeding assays. See, for example Marrone *et al.* (1985) *J. of Economic Entomology* 78:290-293. Such assays can include contacting plants with one or more pests and determining the plant's ability to survive and/or cause the death of the pests. Examples of mutations that result in increased toxicity are found in Schnepf *et al.* (1998) *Microbiol. Mol. Biol. Rev.* 62:775-806.

Alternatively, alterations may be made to the protein sequence of many proteins at the amino or carboxy terminus without substantially affecting activity. This can include insertions, deletions, or alterations introduced by modern molecular methods, such as PCR, including PCR amplifications that alter or extend the protein coding sequence by virtue of inclusion of amino acid encoding sequences in the oligonucleotides utilized in the PCR amplification. Alternatively, the protein sequences added can include

entire protein-coding sequences, such as those used commonly in the art to generate protein fusions. Such fusion proteins are often used to (1) increase expression of a protein of interest (2) introduce a binding domain, enzymatic activity, or epitope to facilitate either protein purification, protein detection, or other experimental uses known in the art (3) target secretion or translation of a protein to a subcellular organelle, such as the periplasmic space of Gram-negative bacteria, or the endoplasmic reticulum of eukaryotic cells, the latter of which often results in glycosylation of the protein.

Variant nucleotide and amino acid sequences of the present invention also encompass sequences derived from mutagenic and recombinogenic procedures such as DNA shuffling. With such a procedure, one or more different delta-endotoxin protein coding regions can be used to create a new delta-endotoxin protein possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*.

For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the delta-endotoxin gene of the invention and other known delta-endotoxin genes to obtain a new gene coding for a protein with an improved property of interest, such as an increased insecticidal activity. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Crameri *et al.* (1997) *Nature Biotech.* 15:436-438; Moore *et al.* (1997) *J. Mol. Biol.* 272:336-347; Zhang *et al.* (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Crameri *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

Domain swapping or shuffling is another mechanism for generating altered delta-endotoxin proteins. Domains II and III may be swapped between delta-endotoxin proteins, resulting in hybrid or chimeric toxins with improved pesticidal activity or target spectrum. Methods for generating recombinant proteins and testing them for pesticidal activity are well known in the art (see, for example, Naimov *et al.* (2001) *Appl. Environ. Microbiol.* 67:5328-5330; de Maagd *et al.* (1996) *Appl. Environ. Microbiol.* 62:1537-1543; Ge *et al.* (1991) *J. Biol. Chem.* 266:17954-17958; Schnepf *et al.* (1990) *J. Biol. Chem.* 265:20923-20930; Rang *et al.* (1999) *Appl. Environ. Microbiol.* 65:2918-2925).

## Plant Transformation

Transformation of plant cells can be accomplished by one of several techniques known in the art. First, one engineers the delta-endotoxin gene in a way that allows its  
5 expression in plant cells. Typically a construct that expresses such a protein would contain a promoter to drive transcription of the gene, as well as a 3' untranslated region to allow transcription termination and polyadenylation. The organization of such constructs is well known in the art. In some instances, it may be useful to engineer the gene such that the resulting peptide is secreted, or otherwise targeted within the plant cell. For  
10 example, the gene can be engineered to contain a signal peptide to facilitate transfer of the peptide to the endoplasmic reticulum. It may also be preferable to engineer the plant expression cassette to contain an intron, such that mRNA processing of the intron is required for expression.

Typically this 'plant expression cassette' will be inserted into a 'plant  
15 transformation vector'. This plant transformation vector may be comprised of one or more DNA vectors needed for achieving plant transformation. For example, it is a common practice in the art to utilize plant transformation vectors that are comprised of more than one contiguous DNA segment. These vectors are often referred to in the art as 'binary vectors'. Binary vectors as well as vectors with helper plasmids are most often  
20 used for *Agrobacterium*-mediated transformation, where the size and complexity of DNA segments needed to achieve efficient transformation is quite large, and it is advantageous to separate functions onto separate DNA molecules. Binary vectors typically contain a plasmid vector that contains the cis-acting sequences required for T-DNA transfer (such as left border and right border), a selectable marker that is engineered to be capable of  
25 expression in a plant cell, and a 'gene of interest' (a gene engineered to be capable of expression in a plant cell for which generation of transgenic plants is desired). Also present on this plasmid vector are sequences required for bacterial replication. The cis-acting sequences are arranged in a fashion to allow efficient transfer into plant cells and expression therein. For example, the selectable marker gene and the gene of interest are  
30 located between the left and right borders. Often a second plasmid vector contains the trans-acting factors that mediate T-DNA transfer from *Agrobacterium* to plant cells. This

plasmid often contains the virulence functions (Vir genes) that allow infection of plant cells by *Agrobacterium*, and transfer of DNA by cleavage at border sequences and vir-mediated DNA transfer, as is understood in the art (Hellens and Mullineaux (2000) *Trends in Plant Science*, 5:446-451). Several types of *Agrobacterium* strains (e.g.

5 LBA4404, GV3101, EHA101, EHA105, etc.) can be used for plant transformation. The second plasmid vector is not necessary for transforming the plants by other methods such as microprojection, microinjection, electroporation, polyethylene glycol, etc.

In general, plant transformation methods involve transferring heterologous DNA into target plant cells (e.g. immature or mature embryos, suspension cultures, undifferentiated callus, protoplasts, etc.), followed by applying a maximum threshold level of appropriate selection (depending on the selectable marker gene) to recover the transformed plant cells from a group of untransformed cell mass. Explants are typically transferred to a fresh supply of the same medium and cultured routinely. Subsequently, the transformed cells are differentiated into shoots after placing on regeneration medium supplemented with a maximum threshold level of selecting agent. The shoots are then transferred to a selective rooting medium for recovering rooted shoot or plantlet. The transgenic plantlet then grows into a mature plant and produces fertile seeds (e.g. Hiei *et al.* (1994) *The Plant Journal* 6: 271-282; Ishida *et al.* (1996) *Nature Biotechnology* 14: 745-750). Explants are typically transferred to a fresh supply of the same medium and cultured routinely. A general description of the techniques and methods for generating transgenic plantlets are found in Ayres and Park, 1994 (*Critical Reviews in Plant Science* 13:219-239) and Bommineni and Jauhar, 1997 (*Maydica* 42:107-120). Since the transformed material contains many cells; both transformed and non-transformed cells are present in any piece of subjected target callus or tissue or group of cells. The ability to kill non-transformed cells and allow transformed cells to proliferate results in transformed plant cultures. Often, the ability to remove non-transformed cells is a limitation to rapid recovery of transformed plant cells and successful generation of transgenic plants.

Generation of transgenic plants may be performed by one of several methods, including but not limited to introduction of heterologous DNA by *Agrobacterium* into plant cells (*Agrobacterium*-mediated transformation), bombardment of plant cells with

heterologous foreign DNA adhered to particles, and various other non-particle direct-mediated methods (e.g. Hiei *et al.* (1994) *The Plant Journal* 6:271-282; Ishida *et al.* (1996) *Nature Biotechnology* 14:745-750; Ayres and Park (1994) *Critical Reviews in Plant Science* 13:219-239; Bommineni and Jauhar (1997) *Maydica* 42:107-120) to transfer DNA.

Transformation protocols as well as protocols for introducing nucleotide sequences into plants may vary depending on the type of plant or plant cell, i.e., monocot or dicot, targeted for transformation. Suitable methods of introducing nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway *et al.* (1986) *Biotechniques* 4:320-334), electroporation (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, *Agrobacterium*-mediated transformation (U.S. Patent No. 5,563,055; U.S. Patent No. 5,981,840), direct gene transfer (Paszkowski *et al.* (1984) *EMBO J.* 3:2717-2722), and ballistic particle acceleration (see, for example, U.S. Patent No. 4,945,050; U.S. Patent No. 5,879,918; U.S. Patent No. 5,886,244; U.S. Patent No. 5,932,782; Tomes *et al.* (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); McCabe *et al.* (1988) *Biotechnology* 6:923-926); aerosol beam transformation (U.S. Published Application No. 20010026941; U.S. Patent No. 4,945,050; International Publication No. WO 91/00915; U.S. Published Application No. 2002015066); and *Lec1* transformation (WO 00/28058). Also see, Weissinger *et al.* (1988) *Ann. Rev. Genet.* 22:421-477; Sanford *et al.* (1987) *Particulate Science and Technology* 5:27-37; Christou *et al.* (1988) *Plant Physiol.* 87:671-674; McCabe *et al.* (1988) *Bio/Technology* 6:923-926; Finer and McMullen (1991) *In Vitro Cell Dev. Biol.* 27P:175-182; Singh *et al.* (1998) *Theor. Appl. Genet.* 96:319-324; Datta *et al.* (1990) *Biotechnology* 8:736-740; Klein *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309; U.S. Patent No. 5,240,855; U.S. Patent Nos. 5,322,783 and 5,324,646; Tomes *et al.* (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg (Springer-Verlag, Berlin); Klein *et al.* (1988) *Plant Physiol.* 91:440-444; Hooykaas-Van Slogteren *et al.* (1984) *Nature (London)* 311:763-764; U.S. Patent No. 5,736,369; Bytebier *et al.* (1987) *Proc. Natl.*

*Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209; Kaeppler *et al.* (1990) *Plant Cell Reports* 9:415-418 and Kaeppler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566; D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505; Li *et al.* (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford (1995) *Annals of Botany* 75:407-413 (rice); Osjoda *et al.* (1996) *Nature Biotechnology* 14:745-750; all of which are herein incorporated by reference.

Following integration of heterologous foreign DNA into plant cells, one then applies a maximum threshold level of appropriate selection in the medium to kill the untransformed cells and separate and proliferate the putatively transformed cells that survive from this selection treatment by transferring regularly to a fresh medium. By continuous passage and challenge with appropriate selection, one identifies and proliferates the cells that are transformed with the plasmid vector. Then molecular and biochemical methods will be used for confirming the presence of the integrated heterologous gene of interest in the genome of transgenic plant.

The cells that have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having constitutive expression of the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. In this manner, the present invention provides transformed seed (also referred to as "transgenic seed") having a nucleotide construct of the invention, for example, an expression cassette of the invention, stably incorporated into their genome.

The delta-endotoxin sequences of the invention may be provided in expression cassettes for expression in the plant of interest. The cassette will include 5' and 3' regulatory sequences operably linked to a sequence of the invention. By "operably linked" is intended a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence

corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame. The cassette may additionally contain at least one additional gene to be cotransformed into the organism. Alternatively,  
5 the additional gene(s) can be provided on multiple expression cassettes.

Such an expression cassette is provided with a plurality of restriction sites for insertion of the delta-endotoxin sequence to be under the transcriptional regulation of the regulatory regions.

The expression cassette will include in the 5'-3' direction of transcription, a  
10 transcriptional and translational initiation region (i.e., a promoter), a DNA sequence of the invention, and a transcriptional and translational termination region (i.e., termination region) functional in plants. The promoter may be native or analogous, or foreign or heterologous, to the plant host and/or to the DNA sequence of the invention. Additionally, the promoter may be the natural sequence or alternatively a synthetic  
15 sequence. Where the promoter is "native" or "homologous" to the plant host, it is intended that the promoter is found in the native plant into which the promoter is introduced. Where the promoter is "foreign" or "heterologous" to the DNA sequence of the invention, it is intended that the promoter is not the native or naturally occurring promoter for the operably linked DNA sequence of the invention.

20 The termination region may be native with the transcriptional initiation region, may be native with the operably-linked DNA sequence of interest, may be native with the plant host, or may be derived from another source (i.e., foreign or heterologous to the promoter, the DNA sequence of interest, the plant host, or any combination thereof). Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such  
25 as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.* (1990) *Gene* 91:151-158; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891-7903; and Joshi *et al.* (1987) *Nucleic Acid Res.* 15:9627-9639.

30 Where appropriate, the gene(s) may be optimized for increased expression in the transformed host cell. That is, the genes can be synthesized using host cell-preferred



codons for improved expression, or may be synthesized using codons at a host-preferred codon usage frequency. Generally, the GC content of the gene will be increased. See, for example, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for a discussion of host-preferred codon usage. Methods are known in the art for synthesizing host-preferred genes. See, for example, U.S. Patent Nos. 6,320,100; 6,075,185; 5,380,831; and 5,436,391, U.S. Published Application Nos. 20040005600 and 20010003849, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference.

In one embodiment, the nucleic acids of interest are targeted to the chloroplast for expression. In this manner, where the nucleic acid of interest is not directly inserted into the chloroplast, the expression cassette will additionally contain a nucleic acid encoding a transit peptide to direct the gene product of interest to the chloroplasts. Such transit peptides are known in the art. See, for example, Von Heijne *et al.* (1991) *Plant Mol. Biol. Rep.* 9:104-126; Clark *et al.* (1989) *J. Biol. Chem.* 264:17544-17550; Della-Cioppa *et al.* (1987) *Plant Physiol.* 84:965-968; Romer *et al.* (1993) *Biochem. Biophys. Res. Commun.* 196:1414-1421; and Shah *et al.* (1986) *Science* 233:478-481.

Methods for transformation of chloroplasts are known in the art. See, for example, Svab *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530; Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab and Maliga (1993) *EMBO J.* 12:601-606. The method relies on particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination. Additionally, plastid transformation can be accomplished by transactivation of a silent plastid-borne transgene by tissue-preferred expression of a nuclear-encoded and plastid-directed RNA polymerase. Such a system has been reported in McBride *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:7301-7305.

The nucleic acids of interest to be targeted to the chloroplast may be optimized for expression in the chloroplast to account for differences in codon usage between the plant nucleus and this organelle. In this manner, the nucleic acids of interest may be synthesized using chloroplast-preferred codons. See, for example, U.S. Patent No. 5,380,831, herein incorporated by reference.

### Evaluation of Plant Transformation

Following introduction of heterologous foreign DNA into plant cells, the transformation or integration of heterologous gene in the plant genome is confirmed by various methods such as analysis of nucleic acids, proteins and metabolites associated with the integrated gene.

PCR Analysis: PCR analysis is a rapid method to screen transformed cells, tissue or shoots for the presence of incorporated gene at the earlier stage before transplanting into the soil (Sambrook and Russell, 2001) PCR is carried out using oligonucleotide primers specific to the gene of interest or *Agrobacterium* vector background, etc.

Southern Analysis: Plant transformation is confirmed by Southern blot analysis of genomic DNA (Sambrook and Russell, 2001). In general, total DNA is extracted from the transformant, digested with appropriate restriction enzymes, fractionated in an agarose gel and transferred to a nitrocellulose or nylon membrane. The membrane or “blot” then is probed with, for example, radiolabeled <sup>32</sup>P target DNA fragment to confirm the integration of introduced gene in the plant genome according to standard techniques (Sambrook and Russell, 2001. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Northern Analysis: RNA is isolated from specific tissues of transformant, fractionated in a formaldehyde agarose gel, blotted onto a nylon filter according to standard procedures that are routinely used in the art (Sambrook, J., and Russell, D.W. 2001. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Expression of RNA encoded by the delta-endotoxin is then tested by hybridizing the filter to a radioactive probe derived from a delta-endotoxin, by methods known in the art (Sambrook and Russell, 2001)

Western blot and Biochemical assays: Western blot and biochemical assays and the like may be carried out on the transgenic plants to confirm the determine the presence of protein encoded by the delta-endotoxin gene by standard procedures (Sambrook, J., and

Russell, D.W. 2001. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) using antibodies that bind to one or more epitopes present on the delta-endotoxin protein.

5     Pesticidal Activity in Plants

           In another aspect of the invention, one may generate transgenic plants expressing delta-endotoxin that have pesticidal activity. Methods described above by way of example may be utilized to generate transgenic plants, but the manner in which the transgenic plant cells are generated is not critical to this invention. Methods known or  
10     described in the art such as *Agrobacterium*-mediated transformation, aerosol beam, biolistic transformation, and non-particle-mediated methods may be used at the discretion of the experimenter. Plants expressing delta-endotoxin may be isolated by common methods described in the art, for example by transformation of callus, selection of transformed callus, and regeneration of fertile plants from such transgenic callus. In such  
15     process, one may use any gene as a selectable marker so long as its expression in plant cells confers ability to identify or select for transformed cells.

           A number of markers have been developed for use with plant cells, such as resistance to chloramphenicol, the aminoglycoside G418, hygromycin, or the like. Other genes that encode a product involved in chloroplast metabolism may also be used as  
20     selectable markers. For example, genes that provide resistance to plant herbicides such as glyphosate, bromoxynil, or imidazolinone may find particular use. Such genes have been reported (Stalker *et al.* (1985) *J. Biol. Chem.* 263:6310-6314 (bromoxynil resistance nitrilase gene); and Sathasivan *et al.* (1990) *Nucl. Acids Res.* 18:2188 (AHAS imidazolinone resistance gene).

25     Fertile plants expressing delta-endotoxin may be tested for pesticidal activity, and the plants showing optimal activity selected for further breeding. Methods are available in the art to assay for pest activity. Generally, the protein is mixed and used in feeding assays. See, for example Marrone *et al.* (1985) *J. of Economic Entomology* 78:290-293.

### Use in Pesticidal Control

General methods for employing the strains of the invention in pesticide control or in engineering other organisms as pesticidal agents are known in the art. See, for example U.S. Patent No. 5,039,523 and EP 0480762A2.

5       The *Bacillus* strains of the invention or the microorganisms that have been genetically altered to contain the pesticidal gene and protein may be used for protecting agricultural crops and products from pests. In one aspect of the invention, whole, i.e., unlysed, cells of a toxin (pesticide)-producing organism are treated with reagents that prolong the activity of the toxin produced in the cell when the cell is applied to the  
10       environment of target pest(s).

Alternatively, the pesticide is produced by introducing a heterologous gene into a cellular host. Expression of the heterologous gene results, directly or indirectly, in the intracellular production and maintenance of the pesticide. In one aspect of this invention, these cells are then treated under conditions that prolong the activity of the toxin  
15       produced in the cell when the cell is applied to the environment of target pest(s). The resulting product retains the toxicity of the toxin. These naturally encapsulated pesticides may then be formulated in accordance with conventional techniques for application to the environment hosting a target pest, e.g., soil, water, and foliage of plants. See, for example EPA 0192319, and the references cited therein. Alternatively, one may  
20       formulate the cells expressing the genes of this invention such as to allow application of the resulting material as a pesticide.

The active ingredients of the present invention are normally applied in the form of compositions and can be applied to the crop area or plant to be treated, simultaneously or in succession, with other compounds. These compounds can be fertilizers, weed killers,  
25       cryoprotectants, surfactants, detergents, pesticidal soaps, dormant oils, polymers, and/or time-release or biodegradable carrier formulations that permit long-term dosing of a target area following a single application of the formulation. They can also be selective herbicides, chemical insecticides, virucides, microbicides, amoebicides, pesticides, fungicides, bacteriocides, nematocides, molluscicides or mixtures of several of these  
30       preparations, if desired, together with further agriculturally acceptable carriers, surfactants or application-promoting adjuvants customarily employed in the art of

formulation. Suitable carriers and adjuvants can be solid or liquid and correspond to the substances ordinarily employed in formulation technology, e.g. natural or regenerated mineral substances, solvents, dispersants, wetting agents, tackifiers, binders or fertilizers. Likewise the formulations may be prepared into edible "baits" or fashioned into pest

5 "traps" to permit feeding or ingestion by a target pest of the pesticidal formulation.

Preferred methods of applying an active ingredient of the present invention or an agrochemical composition of the present invention which contains at least one of the pesticidal proteins produced by the bacterial strains of the present invention are leaf application, seed coating and soil application. The number of applications and the rate of application depend on the intensity of infestation by the corresponding pest.

The composition may be formulated as a powder, dust, pellet, granule, spray, emulsion, colloid, solution, or such like, and may be preparable by such conventional means as desiccation, lyophilization, homogenation, extraction, filtration, centrifugation, sedimentation, or concentration of a culture of cells comprising the polypeptide. In all such compositions that contain at least one such pesticidal polypeptide, the polypeptide may be present in a concentration of from about 1% to about 99% by weight.

Lepidopteran or coleopteran pests may be killed or reduced in numbers in a given area by the methods of the invention, or may be prophylactically applied to an environmental area to prevent infestation by a susceptible pest. Preferably the pest ingests, or is contacted with, a pesticidally-effective amount of the polypeptide. By "pesticidally-effective amount" is intended an amount of the pesticide that is able to bring about death to at least one pest, or to noticeably reduce pest growth, feeding, or normal physiological development. This amount will vary depending on such factors as, for example, the specific target pests to be controlled, the specific environment, location, plant, crop, or agricultural site to be treated, the environmental conditions, and the method, rate, concentration, stability, and quantity of application of the pesticidally-effective polypeptide composition. The formulations may also vary with respect to climatic conditions, environmental considerations, and/or frequency of application and/or severity of pest infestation.

30 The pesticide compositions described may be made by formulating either the bacterial cell, crystal and/or spore suspension, or isolated protein component with the

desired agriculturally-acceptable carrier. The compositions may be formulated prior to administration in an appropriate means such as lyophilized, freeze-dried, desiccated, or in an aqueous carrier, medium or suitable diluent, such as saline or other buffer. The formulated compositions may be in the form of a dust or granular material, or a suspension in oil (vegetable or mineral), or water or oil/water emulsions, or as a wettable powder, or in combination with any other carrier material suitable for agricultural application. Suitable agricultural carriers can be solid or liquid and are well known in the art. The term "agriculturally-acceptable carrier" covers all adjuvants, inert components, dispersants, surfactants, tackifiers, binders, etc. that are ordinarily used in pesticide formulation technology; these are well known to those skilled in pesticide formulation. The formulations may be mixed with one or more solid or liquid adjuvants and prepared by various means, e.g., by homogeneously mixing, blending and/or grinding the pesticidal composition with suitable adjuvants using conventional formulation techniques. Suitable formulations and application methods are described in U.S. Patent No. 6,468,523, herein incorporated by reference.

"Pest" includes but is not limited to, insects, fungi, bacteria, nematodes, mites, ticks, and the like. Insect pests include insects selected from the orders Coleoptera, Diptera, Hymenoptera, Lepidoptera, Mallophaga, Homoptera, Hemiptera, Orthoptera, Thysanoptera, Dermaptera, Isoptera, Anoplura, Siphonaptera, Trichoptera, etc., particularly Coleoptera, Lepidoptera, and Diptera.

Insect pests include insects selected from the orders Coleoptera, Diptera, Hymenoptera, Lepidoptera, Mallophaga, Homoptera, Hemiptera, Orthoptera, Thysanoptera, Dermaptera, Isoptera, Anoplura, Siphonaptera, Trichoptera, etc., particularly Coleoptera and Lepidoptera. Insect pests of the invention for the major crops include: Maize: *Ostrinia nubilalis*, European corn borer; *Agrotis ipsilon*, black cutworm; *Helicoverpa zea*, corn earworm; *Spodoptera frugiperda*, fall armyworm; *Diatraea grandiosella*, southwestern corn borer; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Diatraea saccharalis*, sugarcane borer; *Diabrotica virgifera*, western corn rootworm; *Diabrotica longicornis barberi*, northern corn rootworm; *Diabrotica undecimpunctata howardi*, southern corn rootworm; *Melanotus* spp., wireworms; *Cyclocephala borealis*, northern masked chafer (white grub); *Cyclocephala immaculata*, southern masked chafer

(white grub); *Popillia japonica*, Japanese beetle; *Chaetocnema pulicaria*, corn flea beetle; *Sphenophorus maidis*, maize billbug; *Rhopalosiphum maidis*, corn leaf aphid; *Anuraphis maidiradicis*, corn root aphid; *Blissus leucopterus leucopterus*, chinch bug; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus sanguinipes*, migratory grasshopper; *Hylemya platura*, seedcorn maggot; *Agromyza parvicornis*, corn blot leafminer; *Anaphothrips obscurus*, grass thrips; *Solenopsis milesta*, thief ant; *Tetranychus urticae*, twospotted spider mite; Sorghum: *Chilo partellus*, sorghum borer; *Spodoptera frugiperda*, fall armyworm; *Helicoverpa zea*, corn earworm; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Feltia subterranea*, granulate cutworm; *Phyllophaga crinita*, white grub; *Eleodes*, *Conoderus*, and *Aeolus* spp., wireworms; *Oulema melanopus*, cereal leaf beetle; *Chaetocnema pulicaria*, corn flea beetle; *Sphenophorus maidis*, maize billbug; *Rhopalosiphum maidis*, corn leaf aphid; *Sipha flava*, yellow sugarcane aphid; *Blissus leucopterus leucopterus*, chinch bug; *Contarinia sorghicola*, sorghum midge; *Tetranychus cinnabarinus*, carmine spider mite; *Tetranychus urticae*, twospotted spider mite; Wheat: *Pseudaletia unipunctata*, army worm; *Spodoptera frugiperda*, fall armyworm; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Agrotis orthogonia*, western cutworm; *Elasmopalpus lignosellus*, lesser cornstalk borer; *Oulema melanopus*, cereal leaf beetle; *Hypera punctata*, clover leaf weevil; *Diabrotica undecimpunctata howardi*, southern corn rootworm; Russian wheat aphid; *Schizaphis graminum*, greenbug; *Macrosiphum avenae*, English grain aphid; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus differentialis*, differential grasshopper; *Melanoplus sanguinipes*, migratory grasshopper; *Mayetiola destructor*, Hessian fly; *Sitodiplosis mosellana*, wheat midge; *Meromyza americana*, wheat stem maggot; *Hylemya coarctata*, wheat bulb fly; *Frankliniella fusca*, tobacco thrips; *Cephus cinctus*, wheat stem sawfly; *Aceria tulipae*, wheat curl mite; Sunflower: *Suleima helianthana*, sunflower bud moth; *Homoeosoma electellum*, sunflower moth; *zygogramma exclamationis*, sunflower beetle; *Bothyrus gibbosus*, carrot beetle; *Neolasioptera murtfeldtiana*, sunflower seed midge; Cotton: *Heliothis virescens*, cotton budworm; *Helicoverpa zea*, cotton bollworm; *Spodoptera exigua*, beet armyworm; *Pectinophora gossypiella*, pink bollworm; *Anthonomus grandis*, boll weevil; *Aphis gossypii*, cotton aphid; *Pseudatomoscelis seriatus*, cotton fleahopper; *Trialeurodes abutilonea*,

bandedwinged whitefly; *Lygus lineolaris*, tarnished plant bug; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus differentialis*, differential grasshopper; *Thrips tabaci*, onion thrips; *Frankliniella fusca*, tobacco thrips; *Tetranychus cinnabarinus*, carmine spider mite; *Tetranychus urticae*, twospotted spider mite; Rice: *Diatraea saccharalis*, sugarcane borer; *Spodoptera frugiperda*, fall armyworm; *Helicoverpa zea*, corn earworm; *Colaspis brunnea*, grape colaspis; *Lissorhoptrus oryzophilus*, rice water weevil; *Sitophilus oryzae*, rice weevil; *Nephotettix nigropictus*, rice leafhopper; *Blissus leucopterus leucopterus*, chinch bug; *Acrosternum hilare*, green stink bug; Soybean: *Pseudoplusia includens*, soybean looper; *Anticarsia gemmatilis*, velvetbean caterpillar; *Plathypena scabra*, green cloverworm; *Ostrinia nubilalis*, European corn borer; *Agrotis ipsilon*, black cutworm; *Spodoptera exigua*, beet armyworm; *Heliothis virescens*, cotton budworm; *Helicoverpa zea*, cotton bollworm; *Epilachna varivestis*, Mexican bean beetle; *Myzus persicae*, green peach aphid; *Empoasca fabae*, potato leafhopper; *Acrosternum hilare*, green stink bug; *Melanoplus femurrubrum*, redlegged grasshopper; *Melanoplus differentialis*, differential grasshopper; *Hylemya platura*, seedcorn maggot; *Sericothrips variabilis*, soybean thrips; *Thrips tabaci*, onion thrips; *Tetranychus turkestanii*, strawberry spider mite; *Tetranychus urticae*, twospotted spider mite; Barley: *Ostrinia nubilalis*, European corn borer; *Agrotis ipsilon*, black cutworm; *Schizaphis graminum*, greenbug; *Blissus leucopterus leucopterus*, chinch bug; *Acrosternum hilare*, green stink bug; *Euschistus servus*, brown stink bug; *Delia platura*, seedcorn maggot; *Mayetiola destructor*, Hessian fly; *Petrobia latens*, brown wheat mite; Oil Seed Rape: *Brevicoryne brassicae*, cabbage aphid; *Phyllotreta cruciferae*, Flea beetle; *Mamestra configurata*, Bertha armyworm; *Plutella xylostella*, Diamond-back moth; *Delia* spp., Root maggots.

Nematodes include parasitic nematodes such as root-knot, cyst, and lesion nematodes, including *Heterodera* spp., *Meloidogyne* spp., and *Globodera* spp.; particularly members of the cyst nematodes, including, but not limited to, *Heterodera glycines* (soybean cyst nematode); *Heterodera schachtii* (beet cyst nematode); *Heterodera avenae* (cereal cyst nematode); and *Globodera rostochiensis* and *Globodera pailida* (potato cyst nematodes). Lesion nematodes include *Pratylenchus* spp.

The following examples are offered by way of illustration and not by way of limitation.



## EXPERIMENTAL

### Example 1. Extraction of plasmid DNA

A pure culture of strain ATX13002 was grown in large quantities of rich media.

- 5 The culture was spun to harvest the cell pellet. The cell pellet was then prepared by treatment with SDS allowing breakage of the cell wall and release of DNA. Proteins and large genomic DNA was then precipitated by a high salt concentration. The plasmid DNA was then taken and precipitated by standard ethanol precipitation. The plasmid DNA was separated from any remaining chromosomal DNA by high-speed
- 10 centrifugation through a cesium chloride gradient. The DNA was visualized in the gradient by UV light and the lower plasmid band was extracted using a syringe. This band contained the plasmid DNA from strain ATX13002. Quality of the DNA was checked by visualization on an agarose gel.

### Example 2. Cloning of Genes

The purified plasmid DNA was sheared into 5-10 kb sized fragments and the 5' and 3' single stranded overhangs repaired using T4 DNA polymerase and Klenow fragment in the presence of all four dNTPs, as known in the art. Phosphates were then attached to the 5' ends by treatment with T4 polynucleotide kinase, as known in the art.

- 20 The repaired DNA fragments were then ligated overnight into a standard high copy vector (i.e. pBluescript SK+), suitably prepared to accept the inserts as known in the art (for example by digestion with a restriction enzyme producing blunt ends).

- The quality of the library was analyzed by digesting a subset of clones with a restriction enzyme known to have a cleavage site flanking the cloning site. A high
- 25 percentage of clones were determined to contain inserts, with an average insert size of 5-6 kb.

### Example 3. High Throughput Sequencing of Library Plates

- Once the shotgun library quality was checked and confirmed, colonies were
- 30 grown in a rich broth in 2 ml 96-well blocks overnight at 37°C at a shaking speed of 350

rpm. The blocks were spun to harvest the cells to the bottom of the block. The blocks were then prepared by standard alkaline lysis prep in a high throughput format.

5 The end sequences of clones from this library were then determined for a large number of clones from each block in the following way: The DNA sequence of each clone chosen for analysis was determined using the fluorescent dye terminator sequencing technique (Applied Biosystems) and standard primers flanking each side of the cloning site. Once the reactions had been carried out in the thermocycler, the DNA was precipitated using standard ethanol precipitation. The DNA was resuspended in water and loaded onto a capillary sequencing machine. Each library plate of DNA was  
10 sequenced from either end of the cloning site, yielding two reads per plate over each insert.

#### Example 4. Assembly and Screening of Sequencing Data

- DNA sequences obtained were compiled into an assembly project and aligned  
15 together to form contigs. This can be done efficiently using a computer program, such as Vector NTi, or alternatively by using the Pred/Phrap suite of DNA alignment and analysis programs. These contigs, along with any individual read that may not have been added to a contig, were compared to a compiled database of all classes of known pesticidal genes. Contigs or individual reads identified as having identity to a known  
20 endotoxin or pesticidal gene were analyzed further. A single clone, pAX004, containing DNA showing homology to known endotoxin genes. Therefore, pAX004 was selected for further sequencing.

#### Example 5. Sequencing of pAX004 and Identification of AXMI-004

25 Primers were designed to anneal to pAX004, in a manner such that DNA sequences generated from such primers will overlap existing DNA sequence of this clone(s). This process, known as “oligo walking”, is well known in the art. This process was utilized to determine the entire DNA sequence of the region exhibiting homology to a known endotoxin gene. In the case of pAX004, this process was used to determine the  
30 DNA sequence of the entire clone, resulting in a single nucleotide sequence. The completed DNA sequence was then placed back into the original large assembly for

further validation. This allowed incorporation of more DNA sequence reads into the contig, resulting in 6-7 reads of coverage over the entire region.

Analysis of the DNA sequence of pAX004 by methods known in the art identified an open reading frame with homology to known delta endotoxin genes. This open reading frame is designated as AXMI-004. The DNA sequence of AXMI-004 is provided as SEQ ID NO:1, and the amino acid sequence of the predicted AMXI-004 protein is provided as SEQ ID NO:3. An alternate start site for AXMI-004 at nucleotide 385 of SEQ ID NO:1 generates the amino acid sequence provided as SEQ ID NO:5.

#### Example 6. Homology of AXMI-004 to Known Endotoxin Genes

Searches of DNA and protein databases with the DNA sequence and amino acid sequence of AXMI-004 reveal that AXMI-004 is homologous to known endotoxins. Figure 1 shows an alignment of AXMI-004 with several endotoxins. Blast searches identify cry1Ca as having the strongest block of homology, though the overall sequence identity in the toxic domain is only 43% (see Table 1).

Alignment of AXMI-004 amino acid sequence with the highest scoring proteins identified by blast search.

Table 1. Amino Acid Identity of AXMI-004 with Exemplary Endotoxin Classes

Endotoxin	Percent Amino Acid Identity to AXMI-004	Percent Amino Acid Identity in Toxic Domains
cry1Ac*	17%	30%
cry1Ca*	24%	43%
cry2Aa	12%	12%
cry3Aa	33%	33%
cry1Ia	35%	37%
cry7Aa	19%	31%

#### Example 7. Assays for Pesticidal Activity

The ability of a pesticidal protein to act as a pesticide upon a pest is often assessed in a number of ways. One way well known in the art is to perform a feeding assay. In such a feeding assay, one exposes the pest to a sample containing either compounds to be tested, or control samples. Often this is performed by placing the material to be tested, or a suitable dilution of such material, onto a material that the pest will ingest, such as an artificial diet. The material to be tested may be composed of a liquid, solid, or slurry. The material to be tested may be placed upon the surface and then allowed to dry.

Alternatively, the material to be tested may be mixed with a molten artificial diet, then dispensed into the assay chamber. The assay chamber may be, for example, a cup, a dish, or a well of a microtiter plate.

Assays for sucking pests (for example aphids) may involve separating the test material from the insect by a partition, ideally a portion that can be pierced by the sucking mouthparts of the sucking insect, to allow ingestion of the test material. Often the test material is mixed with a feeding stimulant, such as sucrose, to promote ingestion of the test compound.

Other types of assays can include microinjection of the test material into the mouth, or gut of the pest, as well as development of transgenic plants, followed by test of the ability of the pest to feed upon the transgenic plant. Plant testing may involve isolation of the plant parts normally consumed, for example, small cages attached to a leaf, or isolation of entire plants in cages containing insects.

Other methods and approaches to assay pests are known in the art, and can be found, for example in Robertson, J. L. & H. K. Preisler. 1992. *Pesticide bioassays with arthropods*. CRC, Boca Raton, FL. Alternatively, assays are commonly described in the journals "Arthropod Management Tests" and "Journal of Economic Entomology" or by discussion with members of the Entomological Society of America (ESA).

#### Example 8. Expression of AXMI-004 in *Bacillus*

The 1,890 base pair insecticidal AXMI-004 gene was amplified by PCR from pAX004, and cloned into the *Bacillus* Expression vector pAX916 by methods well known in the art. The resulting clone, pAX920, expressed AXMI-004 protein when

transformed into cells of a *cry(-)* *Bacillus thuringiensis* strain (see Figure 2). The *Bacillus* strain containing pAX920 and expressing the 69 kD AXMI-004 insecticidal protein may be cultured on a variety of conventional growth media. A *Bacillus* strain containing pAX920 was grown in CYS media (10 g/l Bacto-casitone; 3 g/l yeast extract; 6 g/l KH<sub>2</sub>PO<sub>4</sub>; 14 g/l K<sub>2</sub>HPO<sub>4</sub>; 0.5 mM MgSO<sub>4</sub>; 0.05 mM MnCl<sub>2</sub>; 0.05 mM FeSO<sub>4</sub>), until sporulation was evident by microscopic examination. Samples were prepared, and AXMI-004 was tested for insecticidal activity in bioassays against important insect pests.

### *Methods*

To prepare CYS media: 10 g/l Bacto-casitone; 3 g/l yeast extract; 6 g/l KH<sub>2</sub>PO<sub>4</sub>; 14 g/l K<sub>2</sub>HPO<sub>4</sub>; 0.5 mM MgSO<sub>4</sub>; 0.05 mM MnCl<sub>2</sub>; 0.05 mM FeSO<sub>4</sub>. The CYS mix should be pH 7, if adjustment is necessary. NaOH or HCl are preferred. The media is then autoclaved and 100 ml of 10X filtered glucose is added after autoclaving. If the resultant solution is cloudy it can be stirred at room temperature to clear.

### Example 9. N-terminal Amino Acid Sequence of AXMI-004 Expressed in *Bacilli*

Analysis of AXMI-004 expressed in *Bacillus* suggested that the protein product detected in these cultures may be reduced in size relative to the full-length AXMI-004 protein. Since many endotoxin proteins are cleaved at the N-terminus after expression in *Bacillus*, we determined the N-terminus of the AXMI-004 protein resulting from *Bacillus* expression. Protein samples from AXMI-004 were separated on PAGE gels, and the protein transferred to PVDF membrane by methods known in the art. The protein band corresponding to AXMI-004 was excised. The N-terminal amino acid sequence of this protein was determined by serial Edman degradation as known in the art. The sequence obtained was as follows:

ERFDKNDAL

Comparison of this amino acid sequence with the sequence of the full length AXMI-004 (SEQ ID NO:3) demonstrates that this amino sequence results from internal cleavage of the AXMI-004 after expression in *Bacillus*, resulting in a protein with an N-terminus corresponding to amino acid 28 of SEQ ID NO:3 (disclosed as SEQ ID NO:5).

Example 10. Bioassay of AXMI-004 on Insect Pests

Insecticidal activity of AXMI-004 was established utilizing accepted bioassay procedures using a sporulated *Bacillus* cell culture lysate expressing AXMI-004. The *Bacillus* culture was grown in 50 ml CYS media for both standard bioassay and LC<sub>50</sub> bioassays. The cultures were then grown for 2 to 3 days at 30°C, 250 rpm until the cells were sporulated. Sporulation was determined by examining microscopically for the presence of spores. AXMI-004 protein samples were prepared by centrifugation of the sporulated cultures at 12,000 x g for 10 min. The pellet was collected and resuspended in 4 ml 20 mM Tris-HCl, pH 8.0. The suspension was sonicated for 20 seconds (at top power using a micro probe) while placing the tube on ice. The protein concentration of the sample was determined by electrophoresis on an SDS 4-20% gradient acrylamide gel along with a known quantity of bovine serum albumin (BSA) (Figure 2). The concentration of AXMI-004 was determined to be 0.4 µg/ul.

AXMI-004 insecticidal activity was tested using a surface treatment bioassay with artificial diet (Multiple Species diet, Southland Products, Lake Village, Arkansas) prepared as known in the art. Bioassays were carried out by applying the *Bacillus* culture expressing AXMI-004 to the diet surface and allowing the surface to air-dry. Standard bioassays utilized five eggs per well and LC<sub>50</sub> bioassays utilized ten neonate insect larvae per well. The eggs or larvae were applied using a fine tip paintbrush. Standard surface bioassays were carried out in 24 well tissue culture plates. 40ul of each sample was applied to each well. Since each well has a surface area of 2 cm<sup>2</sup> (plate source), a 40 µl cell lysate sample contained approximately 0.4 ug/ul AXMI-004. Bioassays where the LC<sub>50</sub> was determined were done in 48 well tissue culture plates, each well representing a surface area of 1 cm<sup>2</sup> (source) using approximately 20ul of 0.4 µg/ul AXMI-004 per well. The final amount of AXMI-004 protein in each bioassay was approximately 8 µg/cm<sup>2</sup>.

Bioassay trays were sealed with Breathe Easy Sealing Tape (Diversified Biotech, Boston MA). Control samples included media only samples, and wells that were not treated with samples. Bioassays were then held for five days in the dark at 25° C and 65% relative humidity and results recorded.

5

Table 2. Insecticidal Activity of AXMI-004

<u>Insect (Latin Name)</u>	<u>Common Name</u>	<u>Activity of AXMI-004</u>
<i>Ostrinia nubilalis</i>	European Corn Borer	100% mortality
<i>Agrotis ipsilon</i>	Black Cutworm	Stunted
<i>Heliothis zea</i>	Corn Earworm	Stunted
<i>Spodoptera frugiperda</i>	Fall Armyworm	Stunted
<i>Heliothis virescens</i>	Tobacco Budworm	100% mortality
<i>Pectinophora gossypiella</i>	Pink Bollworm	75% mortality
<i>Manduca sexta</i>	Tobacco Hornworm	100% mortality
<i>Trichoplusia ni</i>	Cabbage Looper	100% mortality

AXMI-004 showed strong insecticidal activity (100% mortality) against *Ostrinia nubilalis* and *Heliothis virescens*. AXMI-004 also showed insecticidal activity of 50-75% mortality against *Pectinophora gossypiella*. A concentration of 43 µg/cm<sup>2</sup> AXMI-004 gave 70% mortality against *Pectinophora gossypiella*. AXMI-004 severely stunted the growth of *Agrotis ipsilon*, *Heliothis zea*, and *Spodoptera frugiperda*.

15 Example 11. Quantitation of AXMI-004 Insecticidal Activity Against *Heliothis virescens* and *Ostrinia nubilalis*

The LC<sub>50</sub> of AXMI-004 protein on *Ostrinia nubilalis* and *Heliothis virescens* larvae were determined by testing a range of AXMI-004 protein concentrations in insect bioassays, and applying these protein samples to the surface of insect diet. Mortality was recorded at each protein concentration and analyzed using a Probit analysis program. Results were significant at the 95% confidence interval. Since assays were performed by surface contamination, LC<sub>50</sub>s were determined assuming that the entire protein sample remained at the surface during the assay, with little diffusion below the level ingested by

20

the insects. Thus, the values determined may somewhat underestimate the toxicity of the AXMI-004 protein on the tested insects.

Table 3. LC<sub>50</sub> of AXMI-004 on *Ostrinia nubilalis*

<u>AXMI-004 (µg/ml)</u>	<u># dead/total</u>	<u>% Mortality</u>
1000	40/46	86.9
500	28/45	62.2
250	16/43	32.7
125	12/38	31.6

5

$$LC_{50} = 297 \text{ ng/cm}^2; 95\% \text{ CI} = 218-384.$$

Table 4. LC<sub>50</sub> of AXMI-004 on *Heliothis virescens*

<u>AXMI-004 (µg/ml)</u>	<u># dead/total</u>	<u>% Mortality</u>
8000	35/47	74.5
4000	26/44	59.1
2000	18/42	42.9
1000	6/27	22.2
500	4/36	11.1
250	2/37	5.4

10

$$LC_{50} = 2874 \text{ ng/cm}^2; 95\% \text{ CI} = 2189-3933$$

#### Example 12. Quantitation of AXMI-004 Insecticidal Activity Against *Lygus lineolaris*

Bacterial lysates were prepared by growing the *Bacillus* in 50 ml of CYS media for 60 hours. The *Bacillus* culture was then centrifuged at 12,000 rpm for ten minutes and the supernatant discarded. The pellet was resuspended in 5 ml of 20 mM Tris HCl at pH 8.

15

Bioassays were performed by cutting both the tip and the cap off an Eppendorf tube to form a feeding chamber. The insecticidal protein or control was presented to the insect in a solution that was poured into the cap and covered with parafilm (Pechiney



Plastic Packaging, Chicago IL) that the insect could pierce upon feeding. The Eppendorf tube was placed back on the cap top down and 1<sup>st</sup> or 2<sup>nd</sup> instar *Lygus* nymphs were placed into the Eppendorf chamber with a fine tip brush. The cut Eppendorf tube tip was sealed with parafilm creating an assay chamber. The resultant assay chamber was incubated at ambient temperature cap side down. Insecticidal proteins were tested in a solution of 15% glucose at a concentration of 6.6 ug/ml.

Table 5. Insecticidal Activity of AXMI-004 on *Lygus lineolaris*

Protein	No. Dead/Total	% Mortality
AXMI-004	2/4	50%
Control	0/9	0%

10 Example 13. Vectoring of AXMI-004 for Plant Expression

The AXMI-004 coding region DNA is operably connected with appropriate promoter and terminator sequences for expression in plants. Such sequences are well known in the art and may include the rice actin promoter or maize ubiquitin promoter for expression in monocots, the *Arabidopsis* UBQ3 promoter or CaMV 35S promoter for expression in dicots, and the nos or PinII terminators. Techniques for producing and confirming promoter – gene – terminator constructs also are well known in the art.

The plant expression cassettes described above are combined with an appropriate plant selectable marker to aid in the selections of transformed cells and tissues, and ligated into plant transformation vectors. These may include binary vectors from *Agrobacterium*-mediated transformation or simple plasmid vectors for aerosol or biolistic transformation.

Example 14. Transformation of Maize Cells with AXMI-004

Maize ears are collected 8-12 days after pollination. Embryos are isolated from the ears, and those embryos 0.8-1.5 mm in size are used for transformation. Embryos are plated scutellum side-up on a suitable incubation media, such as DN62A5S media (3.98 g/L N6 Salts; 1 mL/L (of 1000x Stock) N6 Vitamins; 800 mg/L L-Asparagine; 100 mg/L Myo-

inositol; 1.4 g/L L-Proline; 100 mg/L Casaminoacids; 50 g/L sucrose; 1 mL/L (of 1 mg/mL Stock) 2,4-D), and incubated overnight at 25°C in the dark.

The resulting explants are transferred to mesh squares (30-40 per plate), transferred onto osmotic media for 30-45 minutes, then transferred to a beaming plate (see, for example, PCT Publication No. WO/0138514 and U.S. Patent No. 5,240,842).

DNA constructs designed to express AXMI-004 in plant cells are accelerated into plant tissue using an aerosol beam accelerator, using conditions essentially as described in PCT Publication No. WO/0138514. After beaming, embryos are incubated for 30 min on osmotic media, then placed onto incubation media overnight at 25°C in the dark. To avoid unduly damaging beamed explants, they are incubated for at least 24 hours prior to transfer to recovery media. Embryos are then spread onto recovery period media, for 5 days, 25°C in the dark, then transferred to a selection media. Explants are incubated in selection media for up to eight weeks, depending on the nature and characteristics of the particular selection utilized. After the selection period, the resulting callus is transferred to embryo maturation media, until the formation of mature somatic embryos is observed. The resulting mature somatic embryos are then placed under low light, and the process of regeneration is initiated by methods known in the art. The resulting shoots are allowed to root on rooting media, and the resulting plants are transferred to nursery pots and propagated as transgenic plants.

## Materials

### DN62A5S Media

Components	per liter	Source
Chu'S N6 Basal Salt Mixture (Prod. No. C 416)	3.98 g/L	Phytotechnology Labs
Chu's N6 Vitamin Solution (Prod. No. C 149)	1 mL/L (of 1000x Stock)	Phytotechnology Labs
L-Asparagine	800 mg/L	Phytotechnology Labs
Myo-inositol	100 mg/L	Sigma
L-Proline	1.4 g/L	Phytotechnology Labs
Casaminoacids	100 mg/L	Fisher Scientific
Sucrose	50 g/L	Phytotechnology Labs
2,4-D (Prod. No. D-7299)	1 mL/L (of 1 mg/mL Stock)	Sigma

Adjust the pH of the solution to pH to 5.8 with 1N KOH/1N KCl, add Gelrite (Sigma) to 3g/L, and autoclave. After cooling to 50°C, add 2 ml/L of a 5 mg/ml stock solution of Silver Nitrate (Phytotechnology Labs). Recipe yields about 20 plates.

5

Example 15. Transformation of AXMI-004 into Plant Cells by *Agrobacterium*-Mediated Transformation

Ears are collected 8-12 days after pollination. Embryos are isolated from the ears, and those embryos 0.8-1.5 mm in size are used for transformation. Embryos are plated  
10 scutellum side-up on a suitable incubation media, and incubated overnight at 25°C in the dark. However, it is not necessary *per se* to incubate the embryos overnight. Embryos are contacted with an *Agrobacterium* strain containing the appropriate vectors for Ti plasmid mediated transfer for 5-10 min, and then plated onto co-cultivation media for 3 days (25°C in the dark). After co-cultivation, explants are transferred to recovery period media  
15 for five days (at 25°C in the dark). Explants are incubated in selection media for up to eight weeks, depending on the nature and characteristics of the particular selection utilized. After the selection period, the resulting callus is transferred to embryo maturation media, until the formation of mature somatic embryos is observed. The resulting mature somatic embryos are then placed under low light, and the process of  
20 regeneration is initiated as known in the art. The resulting shoots are allowed to root on rooting media, and the resulting plants are transferred to nursery pots and propagated as transgenic plants.

All publications and patent applications mentioned in the specification are indicative of the level of skill of those skilled in the art to which this invention pertains.  
25 All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that  
30 certain changes and modifications may be practiced within the scope of the appended claims.